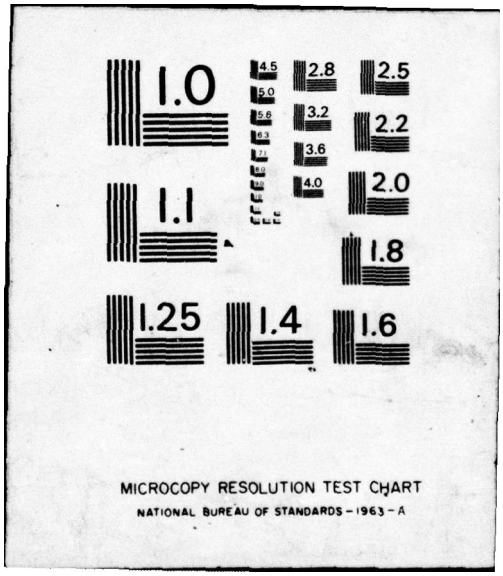
AD-A031 918 PURDUE UNIV LAFAYETTE IND SCHOOL OF ELECTRICAL ENGI--ETC F/G 12/1
EIGENVALUES AND EIGENVECTORS IN SIGNAL ANALYSIS,(U)
SEP 69 K FUKUNAGA, W L KOONTZ N00024-69-C-1265
UNCLASSIFIED TR-EE69-34 NL

END

DATE FILMED
1-77


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
4.5
5.0
5.6
6.3
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS - 1963 - A

TR-EE 69-34

PURDUE UNIVERSITY
SCHOOL OF ELECTRICAL ENGINEERING

EIGENVALUES AND EIGENVECTORS IN SIGNAL ANALYSIS

by
K. Fukunaga
and
W. L. G. Koontz

September 1969
Lafayette, Indiana

Sponsored by
Naval Ship Systems Command, Washington, D. C.
under Contract N00024-69-C-1265

(14) TR-EE 69-34

Figs. 1-5 are not referenced in logical order. Each fig. should be placed nearer to the explanatory paragraphs, e.g. Fig. 1 should be on p. 9 (explanation is on p. 9) Etc.

Purdue University

School of Electrical Engineering

(6) EIGENVALUES AND EIGENVECTORS IN SIGNAL ANALYSIS

by

(10) K. Fukunaga

and

W. L. G. Koontz

(11) September 1969

Lafayette, Indiana

(12) 30 p.

Sponsored By

Naval Ship Systems Command, Washington, D. C.
under Contract (15) NOOO24-69-C-1265

292000 AB

FORWARD

This report is part of a series of reports on operations useful for Automatic Submarine Detection and Classification. Other reports in the series may be obtained by writing Dr. Edward A. Patrick, Associate Professor of Electrical Engineering, Purdue University.

Abstract

The application of the Karhunen-Loeve expansion to signal analysis problems involves eigenvector and eigenvalue calculations. Practical questions which arise are how many eigenvectors should one calculate and how many sample waveforms must one obtain to determine these eigenvectors with a specified accuracy. The answers to these engineering questions are the concern of this paper.

In addition, a rapid algorithm for numerically calculating the eigenvectors and eigenvalues of an autocorrelation matrix is presented which is applicable mainly to signal analysis problems.

1. Introduction

The Karhunen-Loeve expansion is well established as a means of representing a random process in signal analysis. The application of the expansion is conceptually straightforward. We obtain N sample waveforms $x_k(t)$, k=1,...,N, $t\in[0,T]$, and time sample each at n points, t_ℓ, ℓ=1,...,n, $t_\ell\in[0,T]$. We then calculate the autocorrelation matrix, S, whose element, $s_{\ell m}$, is

$$s_{\ell m} = \frac{1}{N} \sum_{k=1}^{N} x_k(t_\ell)\, x_k(t_m), \qquad (1)$$

$$\ell,\ m=1,\ldots,n.$$

Denote the ℓth element of the ith eigenvector of S by $\phi_i(t_\ell)$. Then the Karhunen-Loeve expansion for $x_k(t_\ell)$ is

$$x_k(t_\ell) = \sum_{i=1}^{n} c_{ki}\, \phi_i(t_\ell), \qquad (2)$$

$$k=1,\ldots,N,$$

$$\ell=1,\ldots,n,$$

where

$$c_{ki} = \sum_{\ell=1}^{n} x_k(t_\ell)\, \phi_i(t_\ell), \qquad (3)$$

$$k=1,\ldots,N,$$

$$i=1,\ldots,n.$$

Although the expansion of (2) is an exact representation of $x_k(t_\ell)$'s with n eigenvectors, an approximation can be obtained by selecting a smaller number of eigenvectors. The mean square error of the approximation is the summation of the eigenvalues, λ_i's, whose eigenvectors are not selected.

$$\overline{\epsilon^2} = \frac{1}{N} \sum_{k=1}^{N} \sum_{\ell=1}^{n} \left\{ x_k(t_\ell) - \sum_{i\in I} c_{ki}\, \phi_i(t_\ell) \right\}^2$$

$$= \sum_{i\notin I} \lambda_i. \qquad (4)$$

Thus, the central problem in the Karhunen-Loeve expansions is to calculate

the eigenvectors and eigenvalues and to select the eigenvectors with the larger eigenvalues as the important basis functions to represent a given process. Although this procedure is straightforward, we face engineering problems especially when waveforms are analyzed, mainly because the number of sampling points, n, becomes very large in signal analysis.

This paper is concerned with solving those engineering problems such as

(1) how many samples, N, are needed to estimate the eigenvalues and eigenvectors with a specified accuracy,

(2) how many sampling points, n, should be selected to characterize the process properly.

The quick evaluation of N and n is appreciated especially when n is large such that many eigenvalues and eigenvectors should be calculated.

In addition to the above, this paper offers a program for rapid eigenvector-eigenvalue calculation which has been developed in relation with the above study. The basic approach of this program is to calculate 2n dimensional eigenvalues and eigenvectors from n dimensional eigenvalues and eigenvectors, as the Fast Fourier Transform does for Fourier transforms. The computation time of the program is about three times faster than conventional programs.

2. Eigenvectors and eigenvalues of a perturbed matrix

In this section we derive first order approximations for the eigenvectors and eigenvalues of a perturbed matrix in terms of those of the non-perturbed matrix. These approximations are used in the analysis which follows.

Let S_o be a real, symmetric (nxn) matrix and let dS be a real, symmetric perturbation matrix. Let ϕ_i and λ_i, i=1,...,n, be the eigenvectors and eigenvalues, respectively, of S_o. Assume that the λ_i's are distinct. We wish to obtain a first order approximation of the eigenvectors and eigenvalues of S, in terms of ϕ_i's and λ_i's where

$$S = S_o + dS \tag{5}$$

These may be obtained by retaining the terms of first order or lower of the equation

$$(S_o + dS)(\phi_i + d\phi_i) = (\lambda_i + d\lambda_i)(\phi_i + d\phi_i), \tag{6}$$

where

$$S_o\phi_i = \lambda_i\phi_i. \tag{7}$$

The resulting equation is

$$S_o d\phi_i + dS\phi_i \cong \lambda_i d\phi_i + d\lambda_i\phi_i. \tag{8}$$

To calculate $d\lambda_i$, we premultiply equation (8) by ϕ_i^T and, since $\phi_i^T S_o = \lambda_i\phi_i^T$ and $\phi_i^T\phi_j = \delta_{ij}$, we have

$$d\lambda_i \cong \phi_i^T dS\phi_i. \tag{9}$$

We can write $d\phi_i$ as a linear combination of the ϕ_j's as follows:

$$d\phi_i = \sum_{j=1}^{n} b_{ij}\phi_j, \tag{10}$$

where

$$b_{ij} = \phi_j^T d\phi_i. \tag{11}$$

If we premultiply equation (8) by ϕ_j^T and rearrange, we have for $i \neq j$

$$b_{ij} \cong \frac{\phi_i^T dS \phi_j}{\lambda_i - \lambda_j} . \tag{12}$$

To determine b_{ii} we impose a first order normalization condition on $\phi_i + d\phi_i$, i.e., we require

$$\begin{aligned} ||\phi_i + d\phi_i||^2 &= 1 \\ &\cong ||\phi_i||^2 + 2\, \phi_i^T d\phi_i \\ &= 1 + 2\, \phi_i^T d\phi_i, \end{aligned} \tag{13}$$

and it follows that

$$\phi_i^T d\phi_i = b_{ii} \cong 0. \tag{14}$$

Noting $\phi_i^T S_o \phi_i = \lambda_i$ and $\phi_i^T S_o \phi_j = 0$ for $i \neq j$, we summarize this section as follows:

$$\lambda_i + d\lambda_i \cong \phi_i^T S \phi_i, \tag{15}$$

and

$$b_{ij} \cong \begin{cases} \dfrac{\phi_i^T S \phi_j}{\lambda_i - \lambda_j} & i \neq j \\ 0 & i = j . \end{cases} \tag{16}$$

3. Estimation of eigenvectors and eigenvalues

The estimation of eigenvectors and eigenvalues is made from the sample autocorrelation matrix, S. Therefore, the estimator output is a random variable with its mean and variance. The variance is very much a function of the number of samples, and approaches zero with infinite number of samples. In this section, the variances of estimated eigenvalues and eigenvectors will be discussed. The purpose of this discussion is to find the necessary number of samples to achieve a specified accuracy for the estimation of eigenvectors and eigenvalues.

The autocorrelation matrix, S, is defined by

$$S = E\{XX^T\} , \tag{17}$$

where X is a n dimensional sample vector.

In practice, however, S is not known a priori. Rather, we are given a set of N sample vectors, $X_1,\ldots,X_N$, which we use, as follows, to calculate the sample autocorrelation matrix:

$$\hat{S} = \frac{1}{N} \sum_{k=1}^{N} X_k X_k^T . \tag{18}$$

Let ϕ_i and λ_i, i=1,...,n, be the eigenvectors and eigenvalues respectively of S and similarly define $\hat{\phi}_i$ and $\hat{\lambda}_i$ for $\hat{S}$. Just as $\hat{S}$ is an estimate of S, $\hat{\phi}_i$ and $\hat{\lambda}_i$ are estimates of ϕ_i and λ_i.

The statistics of the eigenvectors and eigenvalues of a matrix of random variables have been studied previously.[1],[2] The general approach is to calculate the distribution of $\hat{S}$ and from this find the distribution of the eigenvectors and eigenvalues.

However, since $\hat{S} \cong S$ for N sufficiently large, we may use the approximations (15) and (16) to express $\hat{\phi}_i$ and $\hat{\lambda}_i$, i.e.,

$$\hat{\phi}_i = \phi_i + \sum_{j=1}^{n} b_{ij}\phi_j, \tag{19}$$

$$\hat{\lambda}_i \cong \phi_i^T \hat{S} \phi_i, \tag{20}$$

$$i=1,\ldots,n,$$

where

$$b_{ij} \cong \begin{cases} \dfrac{\phi_i^T \hat{S} \phi_j}{\lambda_i - \lambda_j} & i \neq j \\ 0 & i=j \ . \end{cases} \tag{21}$$

First, we consider the mean value of the estimates. From (17) and (18),

$$\begin{aligned} E\{\phi_i^T \hat{S} \phi_j\} &= \phi_i^T E\{\hat{S}\} \phi_j \\ &= \phi_i^T S \phi_j \\ &= \lambda_i \delta_{ij} \ . \end{aligned} \tag{22}$$

It follows from (19), (20) and (21) that

$$E\{\hat{\phi}_i\} = \phi_i, \tag{23}$$

and

$$E\{\hat{\lambda}_i\} = \lambda_i. \tag{24}$$

Thus the estimates are seen to be unbiased. The mean square errors of $\hat{\lambda}_i$ and $\hat{\phi}_i$ are given by

$$\begin{aligned} \sigma^2_{\hat{\lambda}_i} &\triangleq E\{(\hat{\lambda}_i - \lambda_i)^2\} = E\{\hat{\lambda}_i^2\} - \lambda_i^2 \\ &\cong E\{(\phi_i^T \hat{S} \phi_i)^2\} - \lambda_i^2 \ , \end{aligned} \tag{25}$$

$$\begin{aligned} \sigma^2_{\hat{\phi}_i} &\triangleq E\{||\hat{\phi}_i - \phi_i||^2\} = \sum_{\substack{j=1 \\ j\neq i}}^{n} E\{b_{ij}^2\} \\ &\cong \sum_{\substack{j=1 \\ j\neq i}}^{n} \frac{E\{(\phi_i^T \hat{S} \phi_j)^2\}}{(\lambda_i - \lambda_j)^2} \ . \end{aligned} \tag{26}$$

The expectations of (25) and (26) can be obtained from the moment generating

functions of the samples, as shown in the Appendix.

(A) Gaussian Cases

For the special case where the X_k's come from a gaussian distribution with mean vector M and autocorrelation matrix S, (25) and (26) become, from (A10) of the Appendix,

$$\sigma^2_{\hat{\lambda}_i} \cong \frac{2}{N}(\lambda_i^2 - \mu_i^4), \tag{27}$$

$$\sigma^2_{\hat{\phi}_i} \cong \frac{1}{N} \sum_{\substack{j=1 \\ j\neq i}}^{n} \frac{\lambda_i\lambda_j - 2\mu_i^2\mu_j^2}{(\lambda_i - \lambda_j)^2}, \tag{28}$$

where

$$\mu_i = \phi_i^T M \qquad (i=1,2,\ldots,n). \tag{29}$$

Equations (27) and (28) indicate that mean square estimation error is the product of 1/N and some coefficient which is independent of N. These coefficients are determined by S and M.

Much simpler expressions are available as upper bounds of $\sigma^2_{\hat{\lambda}_i}$ and $\sigma^2_{\hat{\phi}_i}$ by dropping out the μ terms as follows:

$$\sigma^2_{\hat{\lambda}_i} \leq \frac{2}{N}\lambda_i^2 \quad \text{or} \quad \sigma^2_{\hat{\lambda}_i}/\lambda_i^2 \leq \frac{2}{N}, \tag{30}$$

$$\sigma^2_{\hat{\phi}_i} \leq \frac{1}{N} \sum_{\substack{j=1 \\ j\neq i}}^{n} \frac{\lambda_i\lambda_j}{(\lambda_i - \lambda_j)^2} = \frac{1}{N} \sum_{\substack{j=1 \\ j\neq i}}^{n} \frac{\lambda_j/\lambda_i}{(1 - \lambda_j/\lambda_i)^2} \triangleq \frac{1}{N}\gamma_i. \tag{31}$$

The equalities of (30) and (31) hold when $\mu_i = 0$ for all i's.

Equation (30) shows that $\sigma^2_{\hat{\lambda}_i}/\lambda_i^2$ is simply bounded by 2/N regardless of i and n. On the other hand, from (31), $\sigma^2_{\hat{\phi}_i}$ depends on i and n as well as 1/N. γ_i of (31) is a function of n-1 λ_j/λ_i's (j=1,2,...,n: $j\neq i$). However, γ_i is

roughly determined by a single or a few terms whose λ_j/λ_i's are close to one. This property will be discussed in detail for the following example.

(B) Example: stationary process

Let $x(t)$, $t\in[0,T]$, be a stationary, Gaussian random process with

$$R(\tau) = E\{x(t)x(t-\tau)\}$$
$$= \exp\{-\alpha|\tau|\}. \tag{32}$$

If $x(t)$ is time sampled at $t = \frac{\ell T}{n}$, $\ell=1,\ldots,n$, the S becomes a matrix whose element, $s_{\ell m}$, is

$$s_{\ell m} = \exp\{-\alpha|\ell-m|T/n\},$$
$$= \rho^{|\ell-m|/n}, \tag{33}$$
$$\ell,m=1,\ldots,n,$$

where

$$\rho = \exp\{-\alpha T\}. \tag{34}$$

Suppose that the product αT is such that $\rho \ll 1$. Then the eigenvalues of S are given approximately by

$$\lambda_i \cong K\,\psi(i/T), \tag{35}$$

where $\psi(\cdot)$ is the spectral density of $x(t)$ and K is a constant. Then, since the spectral density of (32) is $\frac{2\alpha}{\omega^2+\alpha^2}$, the ratios of adjacent eigenvalues are approximately

$$\frac{\lambda_{i\pm1}}{\lambda_i} \cong \frac{2\alpha}{\alpha^2+(i\pm1)^2/T^2}\cdot\frac{\alpha^2+i^2/T^2}{2\alpha} \tag{36}$$
$$= 1 - \frac{(1\pm 2i)/\alpha^2T^2}{1+(i\pm1)^2/\alpha^2T^2}.$$

These ratios are independent of n and approach 1 as i becomes large. It can be shown, also, that for large i, $\sigma^2_{\hat{\phi}_i}$ is independent of αT, and hence ρ, and varies essentially as i^2.

The error coefficients, γ_i, i=1,...,n, of (31), for the matrix S of (33) are plotted as a function of n in Figure 1. We see that each γ_i becomes constant as n increases beyond i. Figure 2 shows these steady state values of γ_i as a function of i for various values of ρ. For $i > 4$ we see that γ_i increases linearly with i on the logarithmic plot and is fairly insensitive to ρ. This result is consistent with our approximate theory.

The number of samples required so as to estimate the eigenvectors accurately depends on the number of eigenvectors and hence the dimension. In this example, for a given accuracy the number of samples, N, required increases as n^2.

4. Determining the number of eigenvectors to calculate

In this section, discussion will be given for determining the sampling interval or how many eigenvectors should be calculated for a process in a given time region.

Figure 3(a) shows a waveform x(t). If we sample the waveform at four sampling points t_2, t_4, t_6 and t_8, we have subsequently four eigenvalues as shown in Figure 3(b). Doubling the sampling points from four to eight as shown by the dotted lines in Figure 3(a), we change the previous four eigenvalues and add another four eigenvalues as shown by the dotted lines in Figure 3(b). If four sampling points are enough, the doubling of the sampling points must generate only very small new eigenvalues.

Thus, the basic approach of this paper to check whether we have enough sampling points or not is to calculate the summation of the eigenvalues which are generated by doubling the sampling points.

Suppose the process is sampled at n points in [0,T]. Then the wave vector is

$$X^n = \begin{bmatrix} x(t_2) \\ x(t_4) \\ \vdots \\ x(t_{2n}) \end{bmatrix} \tag{37}$$

and the autocorrelation matrix is

$$\begin{aligned} S^n &= E[X^n X^{n^T}] \\ &= [E\{x(t_{2i})x(t_{2j})\}] \\ &= [R(t_{2i},t_{2j})] , \end{aligned} \tag{38}$$

where $[a_{ij}]$ is a matrix whose ij element is a_{ij}. If the sampling rate is doubled, the wave vector and autocorrelation matrix become respectively

$$X^{2n} = \begin{bmatrix} x(t_2) \\ \vdots \\ x(t_{2n}) \\ x(t_1) \\ \vdots \\ x(t_{2n-1}) \end{bmatrix} = \begin{bmatrix} X^n \\ Y^n \end{bmatrix}, \tag{39}$$

and

$$\begin{aligned} S^{2n} &= E\{X^{2n} X^{2n^T}\} \\ &= \begin{bmatrix} [R(t_{2i}, t_{2j})] & [R(t_{2i}, t_{2j-1})] \\ [R(t_{2i-i}, t_{2j})] & [R(t_{2i-1}, t_{2j-1})] \end{bmatrix} \\ &= \begin{bmatrix} S_{11}^n & S_{12}^n \\ S_{21}^n & S_{22}^n \end{bmatrix} \end{aligned} \tag{40}$$

where S_{11}^n is the same as S^n of (38). The order of $x(t_\ell)$'s is given as (39) only for simplifying discussion. Replacing the order of $x(t_\ell)$'s as $x(t_1)$, $x(t_2)$,... does not change the essential points of the discussion.

If n is large, then S_{11}^n, S_{12}^n, S_{21}^n and S_{22}^n of (40) all become close to each other. Then, S^{2n} can be approximated by

$$S_0^{2n} = \begin{bmatrix} S_{11}^n & S_{11}^n \\ S_{11}^n & S_{11}^n \end{bmatrix}. \tag{41}$$

Assume that the eigenvectors of S_{11}^n, ϕ_i^n, have been calculated. The eigenvector matrix and the eigenvalue matrix of S_0^{2n} are then

$$\Phi^{2n} = \frac{1}{\sqrt{2}} \begin{bmatrix} \Phi^n & \Phi^n \\ \Phi^n & -\Phi^n \end{bmatrix}, \tag{42}$$

$$\Lambda^{2n} = \begin{bmatrix} 2\Lambda^n & 0 \\ 0 & 0 \end{bmatrix}, \tag{43}$$

where

$$\Phi^n = [\phi_1^n \ \dots \ \phi_n^n] \ . \tag{44}$$

By (15), the eigenvalues of S^{2n} can be approximately evaluated by using the eigenvectors of S_0^{2n}, Φ^{2n}. These are the diagonal terms of

$$\begin{aligned} G^{2n} &\simeq \Phi^{2n^T} S^{2n} \Phi^{2n} \\ &= \frac{1}{2} \begin{bmatrix} \Phi^{n^T} & \Phi^{n^T} \\ \Phi^{n^T} & -\Phi^{n^T} \end{bmatrix} \cdot \begin{bmatrix} S_{11}^n & S_{12}^n \\ S_{21}^n & S_{22}^n \end{bmatrix} \cdot \begin{bmatrix} \Phi^n & \Phi^n \\ \Phi^n & -\Phi^n \end{bmatrix} \\ &= \begin{bmatrix} G_{11}^n & G_{12}^n \\ G_{21}^n & G_{22}^n \end{bmatrix}, \end{aligned} \tag{45}$$

where

$$G_{11} = \frac{1}{2} \Phi^{n^T} (S_{11}^n + S_{12}^n + S_{21}^n + S_{22}^n) \Phi^n, \tag{46}$$

$$G_{12} = \frac{1}{2} \Phi^{n^T} (S_{11}^n - S_{12}^n + S_{21}^n - S_{22}^n) \Phi^n, \tag{47}$$

$$G_{21} = \frac{1}{2} \Phi^{n^T} (S_{11}^n + S_{12}^n - S_{21}^n - S_{22}^n) \Phi^n, \tag{48}$$

$$G_{22} = \frac{1}{2} \Phi^{n^T} (S_{11}^n - S_{12}^n - S_{21}^n + S_{22}^n) \Phi^n. \tag{49}$$

If S_{12}^n, S_{21}^n and S_{22}^n are all close to S_{11}^n, G_{11}^n of (46) becomes close to $2\Lambda^n$, and G_{22}^n of (49) becomes close to $\underline{0}$. Thus, the diagonal terms of G_{11}^n are the n eigenvalues which we had for n sampling points and are modified by doubling the sampling points. On the other hand, the diagonal terms of G_{22}^n are the new eigenvalues which are generated by doubling the sampling points.

Therefore, the summation of the new eigenvalues is calculated by

$$\operatorname{tr} G_{22}^n = \sum_{i=1}^{n} \lambda_{2i-1} = \frac{1}{2} \operatorname{tr}(S_{11}^n - S_{12}^n - S_{21}^n + S_{22}^n), \tag{50}$$

where Φ^n can be eliminated from the calculation of trace because $\Phi^{n^T} \Phi^n = \Phi^n \Phi^{n^T} = I$.

If $\sum_{i=1}^{n} \lambda_{2i-1}$ is normalized by $\sum_{i=1}^{n} (\lambda_{2i} + \lambda_{2i-1})$,

$$\beta^n = \frac{\sum_{i=1}^{n} \lambda_{2i-1}}{\sum_{i=1}^{n} (\lambda_{2i}+\lambda_{2i-1})} = \frac{\text{tr } G_{22}^n}{\text{tr}(G_{11}^n+G_{22}^n)} = \frac{\text{tr}(S_{11}^n-S_{12}^n-S_{21}^n+S_{22}^n)}{2\text{tr}(S_{11}^n+S_{22}^n)}. \tag{51}$$

Or, substituting (40) into (51),

$$\beta^n = \frac{1}{2} \frac{\sum_{i=1}^{n} \{R(t_{2i},t_{2i})-R(t_{2i},t_{2i-1})-R(t_{2i-1},t_{2i})+R(t_{2i-1},t_{2i-1})\}}{\sum_{i=1}^{n} \{R(t_{2i},t_{2i})+R(t_{2i-1},t_{2i-1})\}}$$

If $x(t)$ is a stationary process and if $t_k = \frac{kT}{2n}$, then equation (52) reduces to

$$\beta^n = \frac{R(0) - R(\frac{T}{2n})}{2R(0)}. \tag{53}$$

Equation (52) or (53) may be used to determine n by requiring $\beta^n < \epsilon$ where ϵ is fixed by the analyst.

As an example, suppose $R(\tau) = e^{-|t|}$, T=1, and we want to choose n so that $\beta^n < 0.05$. By equation (53)

$$\beta^n = \frac{1 - e^{-\frac{1}{2n}}}{2}$$

$$\simeq \frac{1}{4n} \tag{54}$$

and so we should calculate n=5 eigenvectors.

5. Rapid eigenvector-eigenvalue calculation

In this section, a fast program is offered to calculate eigenvectors and eigenvalues in waveform analysis. For convenience, we shall refer to a program which calculates eigenvectors and eigenvalues as an EIGEN program.

Equations (42) and (43) indicate that 2n dimensional eigenvectors and eigenvalues can be approximated by using n dimensional eigenvectors and eigenvalues, particularly in signal analysis where S_{22}, S_{12} and S_{21} of (40) become close to S_{11} with large n. This is very much analogous to the Fast Fourier Transform, where 2n frequency points can be exactly calculated from n frequency points.

On the other hand, in most EIGEN programs, an iterative procedure is used, starting $\Phi^{(1)}$ and forming a sequence, $\{\Phi^{(n)}\}$, such that

$$\Phi = \lim_{n \to \infty} \Phi^{(n)} , \tag{55}$$

which satisfies

$$\Phi^{T} S \Phi = \Lambda . \tag{56}$$

Furthermore, since a general purpose program can assume little about the structure of S, $\Phi^{(1)}$ is usually taken to be I. Naturally, if we select $\Phi^{(1)}$ which is close to Φ as the starting point, we may reduce the number of iterations.

The basic idea of the Fast EIGEN program of this paper is to combine the above two facts. That is,

(1) approximate 2n dimensional eigenvectors and eigenvalues by n dimensional eigenvectors and eigenvalues, and

(2) use the approximation as $\Phi^{(1)}$.

In most cases where the computation time becomes serious problem, n is large so that the above approximation becomes better.

A better approximation of the 2n dimensional eigenvectors than (42) is

used for the Fast EIGEN program. This is

$$\Phi^{2n} = \frac{1}{\sqrt{2}} \begin{bmatrix} \Phi_1^n & \Phi_1^n \\ \Phi_2^n & -\Phi_2^n \end{bmatrix}, \tag{57}$$

where

$$S_{11}^n \Phi_1^n = \Phi_1^n \Lambda_1 \tag{58}$$

$$S_{22}^n \Phi_2^n = \Phi_2^n \Lambda_2 \tag{59}$$

It is easily seen that Φ^{2n} is an orthogonal matrix. Then,

$$\Phi^{2n^T} S \Phi^{2n} = \begin{bmatrix} \frac{1}{2}(\Lambda_1+\Lambda_2+B+B^T) & \frac{1}{2}(\Lambda_1-\Lambda_2-B+B^T) \\ \frac{1}{2}(\Lambda_1-\Lambda_2+B-B^T) & \frac{1}{2}(\Lambda_1+\Lambda_2-B-B^T) \end{bmatrix}, \tag{60}$$

where

$$B = \Phi_1^T S_{12} \Phi_2 . \tag{61}$$

The approximation of (57) has following two advantages.

(1) Using two sets of n dimensional eigenvectors, Φ_1^n and Φ_2^n, the phase lag information between $x(t_{2i})$-waveforms and $x(t_{2i-1})$ waveforms is taken into account.

(2) In order to calculate $\Phi^{2n^T} S^{2n} \Phi^{2n}$ of (60) only n dimensional matrix multiplication as (61) is required. The number of multiplications for this calculation is roughly estimated as $2n^3$.

When S_{11}, S_{12}, S_{21} and S_{22} are all close to each other with sufficiently large n, then Φ_1 and Φ_2, and Λ_1, Λ_2 and B are also close to each other. Therefore, $\Phi^{2n^T} S \Phi^{2n}$ of (60) is very close to a diagonal matrix. Starting from this matrix and Φ^{2n} of (57), a convensional EIGEN program such as Jacobi's method can start to reduce the off-diagonal terms further [3].

Figure 4 shows how the Fast EIGEN program calculates 2^m dimensional eigenvectors and eigenvalues. Assuming $2^m=8$ for simple explanation, the

procedure is as follows:

(1) Provide eight one dimensional eigenvectors, all 1, and eigenvalues, $s_{\ell\ell}$ (ℓ=1,2,...,8).

(2) Approximate four two dimensional eigenvectors by (57), and calculate (60) for all four cases. Combination should be made as (t_1, t_5), (t_2, t_6), (t_3, t_7) and (t_4, t_8).

(3) Use the results of (60) and (57) as the initial condition and apply the Jacobi's method. Four two dimensional eigenvectors and eigenvalues are obtained.

(4) Repeat (2) and (3), increasing the dimension from two to four, and from four to eight. The number of sets of eigenvectors and eigenvalues are reduced from four to two, and from two to one.

The computation time is roughly estimated as follows: In the Jacobi's method, it has been reported that the computation time, T_J, is

$$T_J = 10\ u\ n^3, \tag{62}$$

where u is a multiplication time and n is the dimension. Assuming that two sets of n/2 eigenvectors and eigenvalues are available, the computation time of one set of n dimensional eigenvectors and eigenvalues, T_n, is

$$T_n = \theta\ T_J = 10\ \theta\ u\ n^3, \tag{63}$$

where θ depends on the initial approximation. For simplicity, let us assume that θ is constant. Then, the computation time of the Fast EIGEN program, T_{FE}, is

$$\begin{aligned} T_{FE} &= 2^0 T_{2^m} + 2^1 T_{2^{m-1}} + \ldots + 2^{m-1} T_{2^1} \\ &= 10\ \theta\ u\ 2^{3m}(1 + 2^{-2} + 2^{-4} + \ldots + 2^{-(2m-1)}) \\ &\cong 10\ \theta\ u\ \frac{2^{3m}}{1-2^{-2}} = \frac{4}{3}\ \theta\ (10\ u\ 2^{3m}) \\ &= \frac{4}{3}\ \theta\ T_J . \end{aligned} \tag{64}$$

Thus, T_{FE}/T_J is almost determined by θ.

For the Jacobi's method, θ is determined experimentally. The computation time is essentially proportional to the number of Jacobi iterations required to achieve convergence to the desired accuracy. Thus by (63) θ is given by

$$\theta = \frac{I_n}{I_J} \tag{65}$$

where I_n is the number of iterations required to obtain Φ^n from $\Phi_1^{n/2}$ and $\Phi_2^{n/2}$ and I_J is the number of iterations required conventionally.[1]

Figure 5 shows curves of I_n and I_J for a matrix S whose i,j element is

$$s_{ij} = \nu^{|i-j|} \tag{66}$$

versus the matrix dimension for various values of ν. The spacing between the Fast EIGEN and conventional curves on the semi-log plot is nearly constant for fixed ν, and θ varies between about 0.25 and 0.35.

Thus computation time can be reduced by the Fast EIGEN approach. The significance of the time saving of course depends on the details of the application.

6. Summary

We have solved in part some of the engineering problems involved with the application of the Karhunen-Loeve expansion. In particular we have outlined procedures for:

1. determining the number of time samples to take from the waveform,

2. determining the number of sample waveforms required to accurately determine the eigenvectors, and

3. reducing the computational effort in numerically determining the eigenvectors.

These procedures in conjunction with a basic understanding of the Karhunen-Loeve expansion allow its straightforward implementation in communication and pattern recognition systems.

Appendix

Calculation of $E\{\phi_i^T \hat{S} \phi_j)^2\}$

Recall that $\hat{S}$ is given by

$$\hat{S} = \frac{1}{N} \sum_{k=1}^{N} X_k X_k^T, \tag{A.1}$$

and thus we have

$$\phi_i^T \hat{S} \phi_j = \frac{1}{N} \sum_{k=1}^{N} Y_i^k Y_j^k, \tag{A.2}$$

where

$$Y_i^k = \phi_i^T X_k \ . \tag{A.3}$$

If both sides of (A.2) are squared and expectation is taken, the result is

$$E\{(\phi_i^T \hat{S} \phi_j)^2\} = \frac{1}{N^2} \sum_{k=1}^{N} \sum_{k'=1}^{N} E[Y_i^k Y_j^k Y_i^{k'} Y_j^{k'}]$$

$$= \frac{1}{N^2} \sum_{\substack{k=1 \\ k \neq k'}}^{N} \sum_{k'=1}^{N} E\{Y_i^k Y_j^k\} \, E\{Y_i^{k'} Y_j^{k'}\}$$

$$+ \frac{1}{N^2} \sum_{k=1}^{N} E\{Y_i^{k^2} Y_j^{k^2}\} \ , \tag{A.4}$$

since the X_k's are independent.

Now

$$E\{Y_i^k Y_j^k\} = \phi_i^T \, E\{X_k X_k^T\} \phi_j$$

$$= \lambda_i \delta_{ij} \quad k=1, \ldots, N. \tag{A.5}$$

and (A.4) may be re-written as

$$E\{(\phi_i^T \hat{S} \phi_j)^2\} = \frac{N-1}{N} \lambda_i^2 \delta_{ij} + \frac{1}{N} E\{Y_i^2 Y_j^2\} \ . \tag{A.6}$$

The supercript on Y is dropped since the X_k's are identically distributed. The last term in (A.6) may be calculated if the moment generating function of X is known. The joint moment generating function of Y_i and Y_j is

$$M_{ij}(t_1,t_2) = E\{\exp(t_1\phi_i^T X + t_2\phi_j^T X)\}$$

$$= E\{\exp[(t_1\phi_i^T + t_2\phi_j^T)X]\}$$

$$= M_X(t_1\phi_i + t_2\phi_j), \tag{A.7}$$

where $M_X(.)$ is the moment generating function of X. It follows then that

$$E\{Y_i^2Y_j^2\} = \begin{cases} \left.\dfrac{\partial^4 M_{ij}(t_1,t_2)}{\partial t_1^2 \partial t_2^2}\right|_{t_1=t_2=0} & i\neq j \\ \left.\dfrac{\partial^4 M_{ij}(t_1,t_2)}{\partial t_1^4}\right|_{t_1=t_2=0} & i=j \end{cases} \tag{A.8}$$

If X comes from a multivariate normal distribution with mean vector M and autocorrelation matrix S, then

$$E\{Y_i^2Y_j^2\} = 2\lambda_i^2\delta_{ij} + \lambda_i\lambda_j - 2\mu_i^2\mu_j^2 . \tag{A.9}$$

Combining (A.9) and (A.6) we have

$$E\{(\phi_i^T\hat{S}\phi_j)^2\} = \lambda_i^2\delta_{ij} + \frac{1}{N}(\lambda_i^2\delta_{ij} + \lambda_i\lambda_j - 2\mu_i^2\mu_j^2) , \tag{A.10}$$

where μ_i is given by (29).

List of References

1. T.W. Anderson, An Introduction to Multivariate Statistical Analysis, New York: John Wiley and Sons, Inc., 1958, pp. 307-325.

2. M.A. Girshick, "On the sampling theory of roots of determinantal equations", Annals of Mathematical Statistics, Vol. 10, 1939, pp. 203-224.

3. J. Greenstadt, "The determination of the characteristic roots of a matrix by the Jacobi method, Mathematical Methods for Digital Computers, A. Ralston and H. S. Wilf, ed., New York: John Wiley and Sons, Inc., 1962, pp. 84-91.

4. P.A. White, "The computation of eigenvalues and eigenvectors of a matrix", J. Soc. Indust. Appl. Math., Vol. 6, 1958, pp. 393-437.

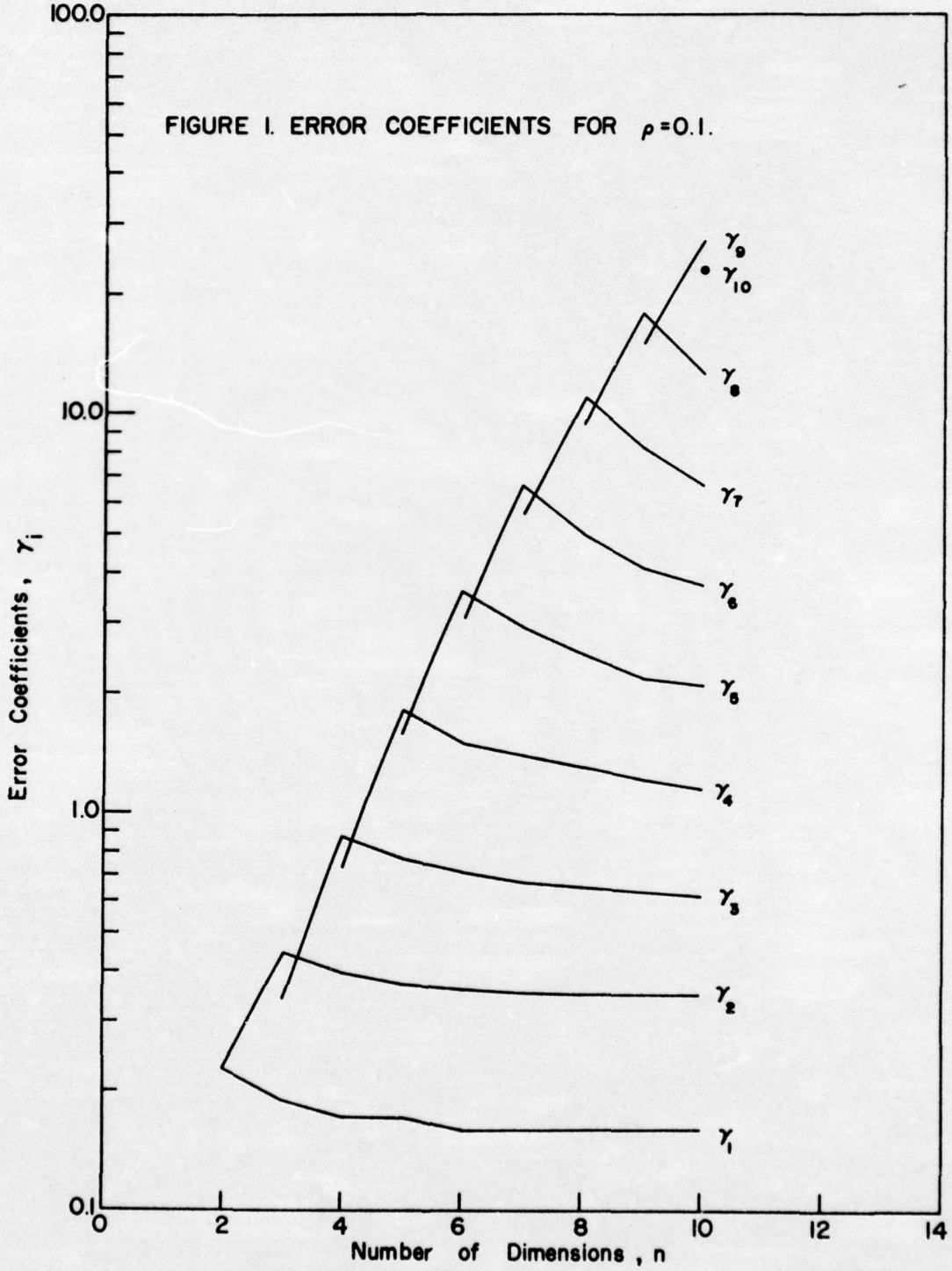


FIGURE 1. ERROR COEFFICIENTS FOR $\rho = 0.1$.

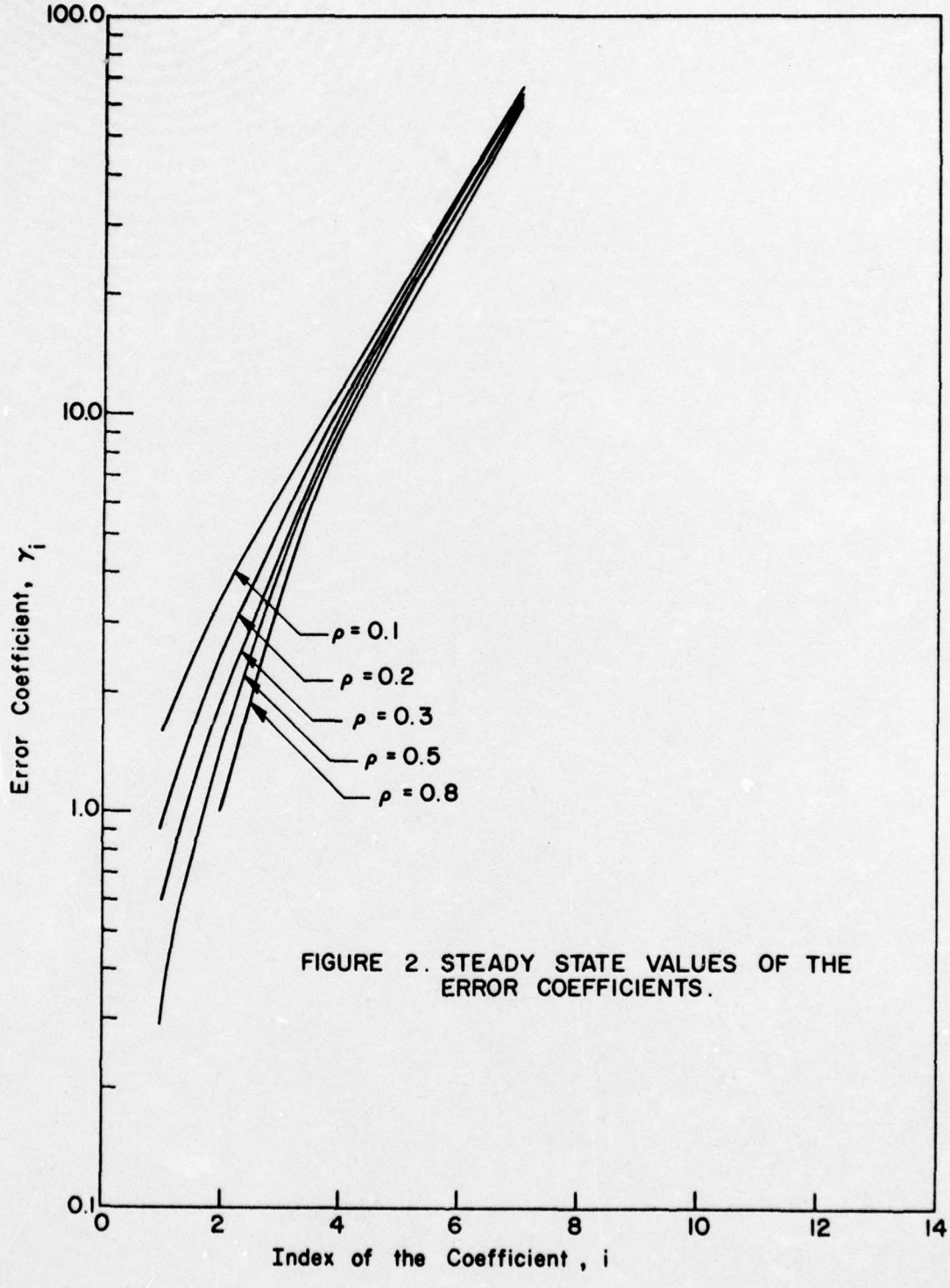


FIGURE 2. STEADY STATE VALUES OF THE ERROR COEFFICIENTS.

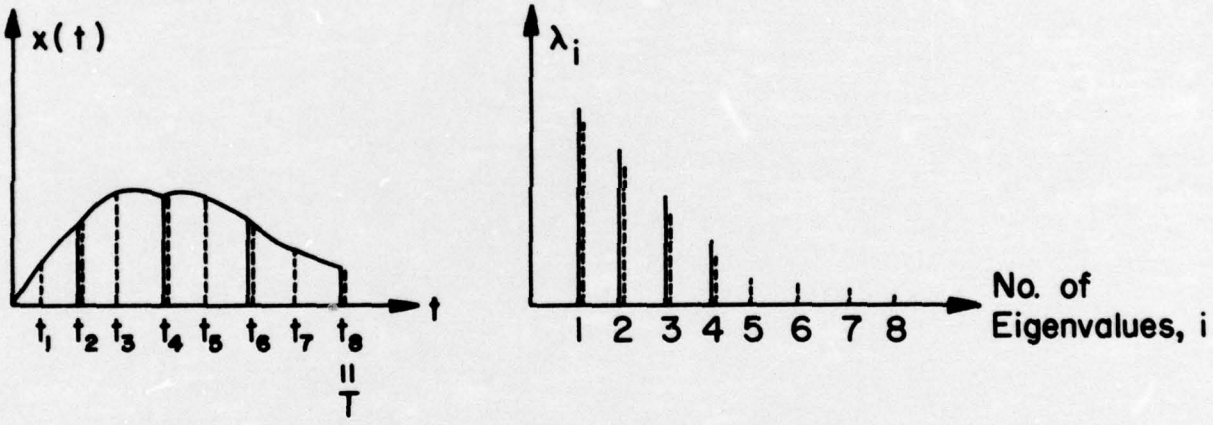


FIGURE 3. A TYPICAL WAVEFORM AND ITS EIGENVALUES.

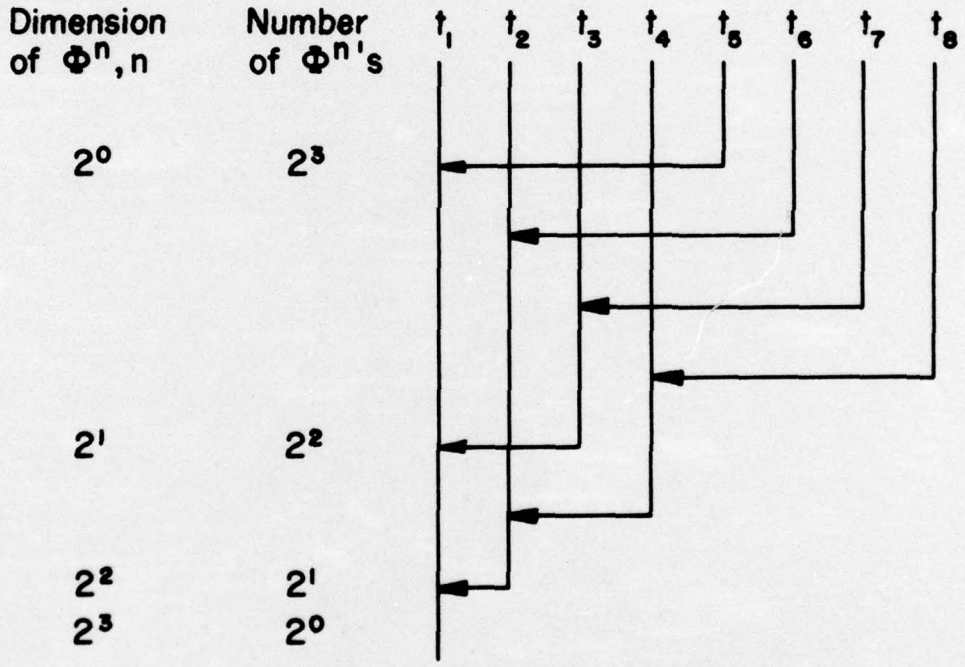


FIGURE 4. DIAGRAM OF THE RULE FOR COMBINING n-DIMENSIONAL EIGENVECTORS TO CALCULATE 2n-DIMENSIONAL EIGENVECTORS.

FIGURE 5. EXPERIMENTAL CURVES FOR DETERMINING THE CONVERGENCE RATIO, θ.